An open-source solution for automatic bug database creation

Péter Gyimesi[a]

[a]Department of Software Engineering, University of Szeged
pgyimesi@inf.u-szeged.hu

Abstract

Several studies about software bugs have used some kind of bug database. These databases usually came from publicly available sources, but some researchers created their own database. Such researches are important, because the more we know about software bugs, the easier it is to prevent or find and fix them. Software developers tend to make mistakes for several reasons: tight deadline, change in the specification, lack of experience, etc. This is the reason why it is substantial to support these kind of researches with bug databases that represent the defects of real software systems well. Depending on the application of the database, it can contain additional information like bug related test cases, static source code metrics, design patterns, process metrics, etc.

During our previous research we constructed a bug database from Java projects on GitHub. This database contains the faulty source code elements (files and classes) and their static source code metrics. Additionally, we used a graph database to compute some process metrics of the buggy source code elements.

In this study, we present our BugHunter tool itself and improve it by using the graph database to locate the buggy source code elements as well. This way our tool is capable of locating the defective source code elements on file, on class, and now even on method level. Our tool uses Neo4j - a popular open-source graph database engine - and its query language, called cypher. The calculations are carried out by running a cypher query, thus our method is highly extensible. To analyze the subject project's source code, we used the free version of the SourceMeter tool. Furthermore, we made our complete tool-chain publicly available as an open-source project on GitHub. This way the complete system can freely be used for studies about software bugs, in particular for studies about bug prediction.
https://github.com/sed-szeged/BugHunterToolchain

Keywords: Bug database, Graph database, Open-source tool

MSC: 68N30